

# **Read Me**

## **Syriac Galen Palimpsest Multispectral Imaging**

Authors: Michael B. Toth  
William A. Christens-Barry  
Elizabeth Ryan  
Kristen St. John  
Doug Emery

**Date: 2 April 2018**

### **1 Syriac Galen Palimpsest Multispectral Imaging**

This data set includes captured and processed data from the multispectral imaging of the Syriac Galen Palimpsest in the Stanford University Library Conservation Department on 15 March, 2018 with a 100 Megapixel (MP) Phase One camera.

The narrowband multispectral imaging system used for this project includes commercial-off-the-shelf hardware and software for digital spectral image capture and viewing with the integrated system. It also includes integrated image processing software for processing and exploitation of the spectral images, utilizing techniques from other cultural heritage studies. The Phase One 100MP medium-format, high-pixel-count camera takes a series of high-quality digital images, each illuminated by a specific wavelength of light. The resulting image set is then digitally processed and combined to reveal residues and features in the object that are not visible to the eye in natural light. These processed images (generated from the captured images) clarify and maximally inform research into residues.

#### **1.1 Camera System**

The camera system utilizes a Phase One iXF Camera System with a 100 MP Achromatic IQ3 CMOS sensor and a 120mm macro lens producing images of over 1200 pixels per inch.

#### **1.2 Illumination System**

The imaging system provides narrowband illumination with light in specific wavelengths from low heat and low maintenance, long-lifetime light emitting diodes (LEDs). It includes two integrated Equipoise Imaging illuminators, each with multiple LEDs, providing illumination for imaging in distinct ultraviolet, visible and infrared spectral bands (see wavelengths in “General File Conventions” below). It is integrated with software to allow simplified system operation and unified metadata capture.

#### **1.3 Filter System**

To capture fluorescence optical emissions from an object, a 6-position motorized filter wheel contains 2-inch square optical glass filters to block reflected light. Filtered images can increase the range of captured information to include both fluorescence

emissions and UV reflectance. This allows the characteristic spectra of substrate, colorant, and contaminant materials to be more completely determined and analyzed. The filter wheel is driven by computer control, with control software and computer interface, with a removable carousel containing a selection of filters (UV bandpass; visible bandpass and longpass filters).

## 1.4 Image Capture Integration

The Spectral XV integrated image capture operating software provides integrated control of the digital camera back, filters and illumination as a single system. This software – based on the CaptureCore application engine developed by Phase One A/S to control camera capture operations and processing workflow – allows streamlined operation and metadata capture from a single interface with simple setup and imaging.

## 1.5 Spectral Imaging Processing

Images are initially processed with ImageJ open-source image processing software and a customized Paleo Toolbox – a spectral imaging toolkit created by Equipoise Imaging, LLC, for applications in cultural heritage imaging. The Paleo toolkit comprises plugin modules that integrate into ImageJ, an open source image processing tool originally developed at the US National Institutes of Health. ImageJ has been widely adopted and extended by scientists working in remote sensing, biological science, and cultural heritage world-wide. It offers a wide range of digital operations for the enhancement and reproduction of non-visible features from the objects based on their spectral response in images captured with the full set of illumination wavelengths and emission bands.

## 2 Rights

All images from Syriac Galen Palimpsest are in the public domain or released under Creative Commons licenses as Free Cultural Works. We request images used in publications be credited to “R.B. Toth Associates/Equipoise Imaging” (this need not be included in data sets).

## 3 Syriac Galen Palimpsest Data Set Contents

The following bifolio of the Syriac Galen Palimpsest were imaged with the 100MP multispectral imaging system in the Stanford University Library Conservation studios:

031r-038v	031v-038r
033r-036v	033v-036r
080r-083v	080v-083r
125r-132v	125v-132r
143r-146v	143v-146r
152r-153v	152v-153r
191r-194v	191v-194r

This data set comprises a core content set of digital images of the Syriac Galen Palimpsest. The data set contains the following folders:

**README.txt file:** This description of the data set in txt form providing an orientation to the data and rights management.

**Output:** Data captured from multispectral imaging of the Syriac Galen Palimpsest that have been converted into TIFF imaged.

**Processed:** Digitally processed images from the captured multispectral images of the Syriac Galen Palimpsest.

The directory structure, starting from the root should be as follows:

Data

SGP100MP

Output

SGP100MP\_SGP-031r038v

SGP100MP\_SGP-031v038r

SGP100MP\_SGP-033r036v

SGP100MP\_SGP-033v036r

SGP100MP\_SGP-080r083v

SGP100MP\_SGP-080v083r

SGP100MP\_SGP-125r132v

SGP100MP\_SGP-125v132r

SGP100MP\_SGP-143r146v

SGP100MP\_SGP-143v146r

SGP100MP\_SGP-152r153v

SGP100MP\_SGP-152v153r

SGP100MP\_SGP-191r194v

SGP100MP\_SGP-191v194r

Processed

SGP100MP\_SGP-033v036r-proc-31Mar2018

SGP100MP\_SGP-143r146v-proc-31Mar2018

SGP100MP\_SGP-143v146r-proc-31Mar2018

SGP100MP\_SGP-191r194v-proc-31Mar2018

SGP100MP\_SGP-191v194r-proc-31Mar2018

SGP100MP\_SGP-031r038v-proc-31Mar2018

SGP100MP\_SGP-031v038r-proc-31Mar2018

SGP100MP\_SGP-033r036v-proc-31Mar2018

StanfordLibrary.json

### 3.1 Core Data

For each bifolio side, the data set provides sequences or stacks of captured and registered images converted to TIFF and JPEG thumbnail images with metadata. These images should be retained as archival images and will be easiest to read with most image viewers.

1. Multispectral images captured using Spectral XV operating software were converted from .IIQ format to 16-bit .TIF format by use of Capture One Software. Converted images have the "\_R" at the end of the rootname.
2. **(If applicable)** Reference "flats" images were smoothed using a gaussian blur operation; smoothed flats images have the string "\_H" at the end of the rootname.
3. IF applicable, subject images that were flattened using reference "\_H" flats images; flattened images have the string "\_F" at the end of the rootname. These flattened images serve as the archival images stored on this site. Not all images are converted and flattened, especially of early imaging that was supplanted by more refined imaging later.

The core data includes:

- Captured Image data consisting of captured IIQ image files converted to TIFF. These are individual images from each of the illuminations taken with different wavelengths.
- Computer Processed images. Images that have been digitally produced through the application of computer algorithms to combine and enhance captured images to enhance visibility of undertext. All processed images are TIFF images or AVI video clips of a series of processed images.
- JSON Files. Metadata is included in associated JSON files for the overall imaging project and each sequence of multispectral images.

Each multispectral capture image folder is provided with descriptive metadata in the JSON file giving details of the image capture for the project, scene and sequence.

This includes basic Archimedes Palimpsest Metadata Standard metadata, such as:

"ProjectID": "100008",

"Name": "SGP 100MegaPixel Phase One",

"Rights": "CC 4.0 BY",

"Publisher": "Owner of the Syriac Galen Palimpsest",

"ProjectNickName": "SGP100MP",

"Creator": "MB Toth",

"Contributors": "R.B. Toth Associates, Stanford University Libraries, Digital Transitions, Phase One A/S, SLAC SSRL",

"Description": "High Resolution Images of Syriac Galen Palimpsest",

"Location": "Stanford University Libraries",

(Images as captured in the Phase One RAW Image file type IIQ format have been stored separately on Syriac Galen Palimpsest drives as backups and working images that can be viewed with Phase One's Capture One software. These RAW images are minimally processed data from the CCD or CMOS sensor of the digital camera.)

## 4 General File Conventions

The unflattened and flattened captured images file names include six fields plus an extension. The initial three fields match the short forms of the project name, scene name, and sequence name. The first and second fields are delimited by "\_", and the second and third fields are delimited by "-". The fourth field consists of a three digit number, indicating the illumination wavelength (in nm), plus a plus a single letter identifier for the camera filter.

The illumination or illuminations used to produce each image cited in the filename of the flattened images include seventeen illumination types, with one symbol each. The illumination symbol is one of the following symbols, or a combination of multiple symbols for processed images:

- 370 - 370 nm UV LED illumination
- 448 - 448 nm visible LED illumination
- 476 - 476 nm visible LED illumination
- 499 - 499 nm visible LED illumination
- 519 - 519 nm visible LED illumination
- 598 - 598 nm visible LED illumination
- 636 - 636 nm visible LED illumination
- 0780 - 780 nm LED IR illumination
- 0870 - 870 nm LED IR illumination
- 0940 - 940 nm LED IR illumination
- N = CLEAR (no filter)
- U = BP365 (UV bandpass filter)
- V = LP400 (long pass filter that passes wavelengths longer than 400 nm - blue and above)
- G = LP515 ((long pass filter that passes wavelengths longer than 515 nm - green and above)
- R = LP515 ((long pass filter that passes wavelengths longer than 590 nm - red and above)
- I = LP715 ((long pass filter that passes wavelengths longer than 715 nm - IR and above)

Examples for unflattened and flattened captured images are respectively:

project\_scene-sequence-<wavelength and filter>\_<index number>\_R.tif and  
project\_scene-sequence-<wavelength and filter>\_<index number>\_F.tif

Processed images amend this naming convention to indicate the type of processing employed. The initials of the individual who created the processed images are (optionally) given in the fourth field of the filename of processed files. Since processing operations most often utilize all of the captured images of a sequence, identification of individual images used as inputs for processing operations are generally omitted. One or more following, underscore-delimited fields describe the processing operations and parameters that were used, appended in order of their application. Within an underscore-delimited field, single hyphens are used to delimit parameter values or image indices used during that processing operation. Usually the parameters refer to the index number of a component image.

A typical filename exemplifies the naming practices used for processed images:

SGP100MP\_SGP-[Bifolio Sequencename]\_WCB\_PCA\_pc05-pc06-pc07.tif

Project name: SGP100MP  
 Scene name: SGP  
 Sequence name: [Bifolio]  
 Creator: William Christens-Barry (this field sometimes not used)  
 Processing: 1. Principal Components Analysis (PCA)  
 2. PCA components 05, 06, and 07 were used in the R, G, and B channels, respectively, of the final (synthetic) RGB TIFF image

In some case, two rounds of PCA processing were performed. Selected components from the first round of PCA processing were used during the second round of processing. In these cases, the string "PCAx2", followed by "-" delimited indices of the first round components that were used in the second round. For example, the file name:

SGP100MP\_SGP-[Bifolio Sequencename]\_WCB\_PCAx2-05-08-12-15\_pc02-pc03--pc06\_RGB.tif

This indicates that William Christens-Barry used components 2, 5, 8, 12, and 15 from the first PCA round in the second round, and used components 02, 03, and 09 from the second round of processing in the red, green, and blue channels, respectively, of the final RGB TIFF image.

Note that the use of "--" as a delimiter indicates that a range of component images was used, e.g. "3--6" would indicate that components 3, 4, 5, and 6 were used. The practice of including a leading "0" is not followed consistently, and that the use of "pc" in the front of a principal component used in an RGB channel is not followed consistently due to excessively long filenames.

Other strings in processed file names include:

"dS8\_BasicRGB" an RGB image has been synthesized from flattened images and desaturated computationally by a factor of x0.8;

"8gs" a single channel (grayscale) image stored in 8-bit format

"desat" an RGB image has been synthesized from flattened images and desaturated visually by an unspecified factor

"Combi" multiple grayscale images (captured or processed) were use combinatorally to create many different synthetic RGB images. The resulting files are very large, and are frequently stored in the "AVI" movie format for viewing and selection of the best images.

The remainder of the file name, including the extension, indicates the file type. These are usually:

1. TIFF still image files, ending in 'tif',
2. JPEG still image files ending in 'jpg'
3. AVI moving image files ending in 'avi'